OPEN ACCESS

University of BRISTOL

Harris, R., & Johnston, R. (2018). Measuring and modelling segregation - New concepts, new methods and new data. *Environment and Planning B: Urban Analytics and City Science*, *45*(6), 999-1002. https://doi.org/10.1177/2399808318808889

Peer reviewed version

Link to published version (if available):
10.1177/2399808318808889

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Sage at https://journals.sagepub.com/doi/abs/10.1177/2399808318808889 . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research
General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

**Editorial: Measuring and modelling segregation – new concepts, new methods, new data**

"Why isn't there just one measure of segregation that everyone agrees upon?"

It's a good question, and one that came from the mouth of an exasperated social commentator who was gathering information about social and ethnic segregation in the UK. The person was seeking a definitive measure of segregation to determine whether it is increasing or decreasing across the country. He was frustrated by the range of choices and would probably have been more so had he had access to this journal and discovered we are about to offer more.

The measurement of segregation has been debated in the social sciences for well over half a century and that debate continues in the pages of this journal and elsewhere. The issue of measurement appears as enduring as segregation itself, and regularly is revisited as new sources of data, new understandings of segregation and its causes, and new computational tools become available to assess its prevalence. Some of the methods measure differences (between the observed amount of a population group and the observed amount of another, for example: the Index of Dissimilarity); some are probabilistic (for instance, the probability that within a randomly selected neighbourhood a randomly selected person is black: the Index of Isolation); some are aspatial and pay no heed to the specific geographical pattern of segregation within the study region (both the aforementioned indices); some are spatial and therefore change with the pattern even if the overall amount of segregation is the same (see, inter alia, Hong et al., 2014); some are better suited for measuring multigroup segregation (the entropy index): most operate at a single scale of analysis but some conceive and measure segregation as a multiscale phenomenon. Many of these approaches are discussed in the papers that follow but for a brief review see Chapter 5 of Kaplan (2018).

The different measures have different mathematical properties that may be more or less desirable according to the context (see, for example, the debate around the so-called Gorard index used in educational research: Allen & Vignoles 2007; Gorard 2007; Johnston & Jones, 2007). However, the underlying differences are not simply mathematical, they are conceptual: they rest on different understandings of what is meant by segregation. As Simpson and Peach (2009, 1379) note, the act of measurement is not neutral but "depends upon what the measurer conceives segregation to be."

It its academic usage, and especially from a geographical perspective, segregation has often been shorthand for residential segregation, which means, in a loose sense, that different population groups are living – to a greater or lesser extent – in different places from one another. In practice, the vagueness of this definition undermines its meaning. For example, it is entirely possible for a population group to be 'segregated' but not at all isolated from other population groups (so segregated but not?). This would occur if the group was small, overwhelmingly concentrated within just a few specific neighbourhoods yet formed a sufficiently small proportion of those neighbourhoods' populations that the neighbourhoods remained diverse. The Index of Dissimilarity would indicate segregation; an Index of Exposure to other groups would not. The concentration might link to spatial clustering – meaning that the neighbourhoods in which the group is concentrated are found near each other – but it need not be so: a group can be concentrated yet also dispersed (a

chess or chequers board has that pattern). Segregation might be a reference to uneven population distributions, spatial concentrations, spatial clustering and/or one group's isolation from others, amongst other possibilities. Those components of segregation are often correlated but they are not simply interchangeable.

Even as one moves from a concept of segregation to a mathematical method that embodies that conception there is still the matter of data to contend with. Commonly segregation is perceived at a sub-regional scale: as something that takes places within cities (traditionally, but also within suburban and rural localities) and between neighbourhoods. To operationalise this requires data about neighbourhoods and the people they contain. Historically the only source that contains the necessary spatial breadth and detail is a national census. But these are infrequently updated – typically even ten years – which creates something of an information vacuum when, as in the UK, there is concern about the persistence of (especially, ethnic) segregation and the concern it may be increasing (that impression is evident in two recent Government reports: *The Casey Review* and the *Integrated Communities Strategy White Paper*: Casey 2016, HM Government 2018). It probably is not increasing – or, if it is, where it is, is the exception rather than the norm – but the paucity of official population data in the seven years since the last Census was made makes it hard to refute the more lurid media headlines such as the following: '*Ghetto Britain: Entire districts segregated, warns report'* (Daily Mail online, November 3, 2016). Censuses also limit the studies to residential segregation which is a starting point for understanding population mixing but omits the various other (and potentially more meaningful) places where people from different social and ethnic backgrounds can come together and mix - schools, the workplace, social societies and so forth.

In short, the answer to the question "why isn't there just one measure of segregation?" is because there cannot be, and any quest to produce a definitive measure fundamentally is misconceived. Instead, the overlaying of various approaches helps to extract a more meaningful understanding of the data and ideally too of the processes that create and the outcomes that arise from segregation. Reardon et al. (2008) write that the study of racial segregation has three main analytical aims: to investigate the patterns of segregation, to investigate the causes of segregation, and to investigate the consequences of segregation. The first of these sheds light on the other two – patterns suggest processes. That light may be partial and much may remain in shadow. Nevertheless, as new understandings, new conceptions, new data, and new methods arise, it is right to revisit the ways in which those patterns of segregation are investigated. In particular, rather than seeing segregation as something that can be captured by a single index, at a single scale of analysis, using a single source of data, increasingly researchers think of segregation as multi-faceted: a 'bundle' of spatially contingent processes and outcomes that operate at a range of scales that are both affected by and, reciprocally, contribute to the carving out of a spatially differentiated social landscape where people come together in some places but not in others – perhaps for reasons of their own choosing (or sense of collective wellbeing and/or identity) and sometimes as the result of spatial inequalities (in earnings, in housing, in education, and so forth). The 'bundle' is what Galster and Shevky (2017) theorise as the differential effects of the spatial opportunity structure that create the spatial foundations of inequality (and of segregation).

Moving from a naïve, 'single-index' view of segregation to a more nuanced approach requires and have been enabled by new data – new sources of consumer data for example (Lan et al. 2018) and various forms of detailed administrative data (e.g. data about state pupils in schools in England; Harris and Johnston 2019) – as well as computational statistical infrastructures that allow for more sophisticated modelling of complex data – to disentangle 'true' evidence of segregation from what might just be randomness (due to the uneven population group sizes in the overall population, for example) (Leckie et al. 2012 and subsequent research). Efforts at harmonising data and most especially their output geography help to facilitate understanding of longer-term trends (Logan et al. 2014; Lloyd et al. 2017).

In soliciting papers for this themed issue of *Environment and Planning B: Urban Analytics and City Science*, we noted that after decades in which studies of residential segregation have been dominated by the use of descriptive indices – such as those of dissimilarity and isolation – there has been a recent surge of interest in developing on those measures to provide greater insights into the observed patterns plus the processes which produce them. One major focus of this work, for example, has been on the multi-scale decision-making processes that underpin the selection of residential areas, and on how the resulting patterns can be quantitatively and visually represented. We invited the authors to showcase their work in this area and are extremely grateful for their response – for the high quality of the work, which is at the cutting-edge of measuring and modelling segregation, and their willingness to share their innovations with a wider audience.

All research generates new challenges and in our chosen field of segregation studies three areas come to mind as ripe for further development. The first is in visualisation: how, for example, to best map the results of a multi-scale analysis and to capture how the patterns of segregation vary at the micro-, meso- and macro-scales? The second is linked to the first: how best to communicate the findings. The methods presented here as elsewhere are sophisticated and sometimes computationally or statistically complex. From experience, there are policy makers who struggle with the existing 'simple' indices that are therefore open to misinterpretation. Communicating the rationale for and the results from more complex models is neither trivial nor unimportant because the numbers they generate have the power to frame the discourse within which segregation is discussed and debated, and to challenge misperceptions where they arise. Third, the link to outcomes – why segregation matters, for whom, where and how the impact may vary with context.

Richard Harris and Ron Johnston
University of Bristol

**References**

Allen R, Vignoles A, 2007. What should an index of school segregation measure? *Oxford Review of Index*, 33(5), 643–668.

Casey L, 2016. *The Casey Review: a review into opportunity and integration.* London: Ministry of Housing, Communities & Local Government.

Galster G, Shevky P, 2017. Spatial Foundations of Inequality: A Conceptual Model and Theoretical Framework. *The Russell Sage Journal of the Social Sciences*, 3(2), 1–33.

Gorard S, 2007. What *does* an index of school segregation measure? A commentary on Allen and Vignoles. *Oxford Review of Index*, 33(5), 669–677.

Harris R, Johnston R, 2019. *Ethnic segregation between schools: is it increasing or decreasing in England?* Bristol: Bristol Press.

HM Government, 2018. *Integrated Communities Strategy Green Paper*. London: Ministry of Housing, Communities & Local Government.

Hong SY, O'Sullivan D, Sadahiro Y, 2014. Implementing Spatial Segregation Measures in R. *PLOS One*, https://doi.org/10.1371/journal.pone.0113767

Johnston R, Jones K, 2010. Measuring Segregation—A Cautionary Tale. *Environment and Planning B*, 42(6), 1264–1270.

Kaplan DH, 2018. *Navigating Ethnicity. Segregation, Placemaking and Difference*. Lanham, MD: Rowman & Littlefield.

Lan T, Kandt J, Longley P, 2018. Ethnicity and Residential Segregation, in P Longley, J Cheshire, A Singleton (eds.) *Consumer Data Research*. UCL Press: London, 71–84.

Leckie GB, Pillinger RJ, Jones K, Goldstein H, 2012. Multilevel modelling of social segregation. Journal of Educational and Behavioral Statistics, 37(1), 3–30.

Lloyd CD, Catney G, Williamson P, Bearman N, 2017. Population Change in Britain. PopChange Briefing 1. Liverpool: Centre for Spatial Demographics Research, University of Liverpool.

Logan JR, Zengwang X, Stults BJ, 2014. Interpolating US Decennial Census Tract Data from as Early as 1970 to 2010: A Longitudinal Tract Database. *The Professional Geographer*, 66(3), 412–420.

Reardon SF, Matthews SA, O'Sullivan D, Lee BA, Firebaugh G, Farrell CR, Bischoff K, 2008. The Geographic Scale of Metropolitan Segregation. *Demography*, 45(3), 489–514.

Simpson L, Peach C, 2009. Measurement and Analysis of Segregation Integration and Diversity: Editorial Introduction. *Journal of Ethnic and Migration Studies*, 35(9), 1377–1380.